DRAFT

Simulation, Test, and Evaluation Process STEP Guidelines

December 2, 1996

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1 Introduction

Purpose

The purpose of this document is to provide a definition for the Simulation, Test, and Evaluation Process (STEP), a description of the process and of the underlying STEP concept, and guidelines for Program Managers, Test & Evaluation Master Plan (TEMP) writers and reviewers, and other users on the application of STEP to an acquisition program.

Background

Under Secretary of Defense for Acquisition and Technology, The Honorable Paul G. Kaminski, in an address to the International Test and Evaluation Associate Symposium in Huntsville, Alabama on October 3, 1995 announced he was "requiring that the Simulation, Test, and Evaluation Process - let's call it STEP - shall be an integral part of our Test and Evaluation Master Plans." In addition, he said "this means our underlying approach will be to model first, simulate, then test, and then iterate the test results back into the model." His "intent is to ensure modeling and simulation (M&S) truly becomes an integral part of our test and evaluation (T&E) planning."

The STEP concept, process, and guidelines presented in this document focus existing processes into a new and more effective approach. They are a synthesis of current Department of Defense (DoD) directives, instructions, standards, and other resources, and of best practices throughout the DoD and industry.

Overview

Chapters 2 through 4 contain the essential elements necessary to understand and apply STEP. The STEP is defined and described in Chapter 2. Chapter 3 presents an approach to developing a strategy for its implementation within an acquisition program while Chapter 4 focuses on how STEP fits within the structure of the TEMP.

Subsequent chapters provide additional details for those users desiring more information in specific areas. Chapter 5 provides information on types of STEP tools which are essential to the process. Chapter 6 discusses application of those tools within STEP throughout the various DoD system acquisition phases. Finally, Chapter 7 discusses how to gain the essential credibility for those tools. Acronyms and definitions used in this document are listed in Appendix A and references are listed in Appendix B.

This page intentionally left blank.

Chapter 2 STEP Definition and Concept

2.1 Definition and Overview

	STEP is an iterative process that integrates both simulation and test for the purpose of evaluating the performance, military worth, or effectiveness of systems to be acquired.

In STEP, simulation and test are integrated, each one depending on the other to be effective and efficient. Simulations provide predictions of the system's performance and effectiveness, while tests are "part of a strategy to provide information regarding risk and risk mitigation, to provide empirical data to validate models and simulations, to permit an assessment of the attainment of technical performance specification and system maturity, and to determine whether systems are operationally effective, suitable, and survivable for intended use." A byproduct of this process is a set of models and simulations with a known degree of credibility providing the potential for reuse in other efforts.

STEP is driven by mission need and system requirements. The product of STEP is information. The information supports acquisition program decisions regarding technical risk, performance, system maturity, operational effectiveness, suitability, and survivability. As such, STEP is a sub-process of the overall acquisition process. STEP applies to all acquisition programs, especially Major Defense Acquisition Programs (MDAP) and Major Automated Information Systems (MAIS).

STEP takes operational and system requirements as inputs, and produces information as output. This information provides the decision maker a sense of how well the system is meeting the operational and system requirements, and how well the risks are being managed. It may provide information on the need to refine system requirements and specifications. Central to the process is evaluation, both for live testing and simulation. Analysis is essential prior to testing to determine what is to be evaluated, during testing to support those activities, and following the testing to extract the information from those activities.

The implications of STEP to the acquisition community are numerous. For example, decision makers can use the predictions of system performance to assess the military worth of the system before any physical prototypes are built. In addition, the development and T&E communities can use predictive simulations to expand the scope of system testing and evaluate system performance in areas that are not readily testable. Furthermore, the developmental and operational T&E communities can use STEP as a mechanism for sharing simulation, test, and evaluation resources and treating developmental and operational testing (DT and OT respectively) in an efficient continuum.

The key ingredient to successfully accomplishing STEP is the development of an interdependent and consistent set of STEP tools. They can be categorized into three types: analysis tools, M&S tools, and testing tools. Analysis tools allow the analyst and tester to focus on that which is essential to evaluate, to monitor the activities as they occur, and to consolidate and analyze the results of their activities. M&S tools include a set of models that together describe the system characteristics and performance at all levels. They are used in a variety of ways from measuring compliance to design requirements through predicting system performance in an operational environment. Figure 1 illustrates the different levels of models and simulations required, and depict how they form a hierarchy. Finally, testing tools include simulations, stimulators, and laboratory facilities that have supported testing for many years.

Throughout STEP, tests are conducted to collect data for evaluating the system and refining and validating the models. Through the model-test-model iteration approach, the set of models mature, culminating in accurate representations of the system with appropriate fidelity which can be used to predict system performance, operational effectiveness and suitability, and to support the acquisition and potentially the training communities. What is new with STEP is the interdependent way in which these tools are applied in support of the acquisition process.

Figure 1. Hierarchy of Modeling and Simulation

2.2 Concept

Figure 2 depicts STEP and its role in the acquisition process. The five areas where it is integrally involved are: (1) as requirements and system specifications are refined; (2) in design development; (3) in support of testing; (4) in providing a link between the developing system and its operational environment, permitting a recurring assessments of its military worth; and (5) after fielding, as a means of evaluating deficiencies and new requirements.

Figure 2. The STEP Concept

In Conjunction with Systems Engineering

STEP begins with the mission need statement and continues throughout the acquisition life cycle. Initially, the mission need and top-level operational requirements are used to develop alternative concepts for the system and to select or develop top-level digital models of the system. For a weapon system program, these top-level models are typically evaluated in theater/campaign and/ or mission/battle-level simulations.

Mission/battle-level models are used to evaluate the ability of a multiple platform force package to perform a specific mission. During this level of analysis, the system models are varied to make trade-offs of the weapon system characteristics, and the simulations are used to predict the performance of the system, ideally under the conditions defined in the system's test scenarios. Evaluations at this time provide a basis for assessing the relative merits of alternative concepts and to support the conceptual design review. The results of these early simulations can also be used to aid in Operational Assessments (OAs) of the system's performance and suitability.

Throughout the acquisition process, the system and functional requirements are refined. They are used for developing the system design and maturing the interdependent models of the system. Still early in the acquisition process, the engagement-level models are employed to predict the system’s performance and to evaluate the system’s effectiveness against appropriate threats.

Mission and functional requirements continue to be refined, and the system reaches the preliminary design stage. As the design detail increases, detailed engineering-level models are built by the system developer as engineering development tools. At this point, many engineering level models are based on phenomenology (science-based) models. Normally, live testing cannot be done up to this point as components, subsystems, and systems have not been built. However, throughout this entire process, M&S tools can be employed to support an evaluation of the alternative designs for their adherence to system requirements. These results can support the decision makers in better managing risk and in refining their requirements. The results of more detailed models and simulations should also be used to refine the higher-level models, which can be used in turn for the reevaluation of the military worth of the system. As a result of this process, a set of models have been created that describe the system characteristics and performance at all levels.

In Conjunction with Design Development

From this time and throughout the remainder of the development cycle, M&S is used both as a predictive tool and with test in an iterative process to evaluate the system design. Both test processes and test facilities can be simulated to provide effective evaluation of design changes. Through this iterative process, the consequences of design changes are evaluated and help translate the most promising design approach into a stable, interoperable, and cost effective design; demonstrate the system's capabilities; and support the preliminary design review.

In Conjunction with Testing

As real hardware and software representations/mock-ups of the system's components or subsystems (brass boards, breadboards) are built, these representations are tested in the laboratory environment (i.e., they replace their digital counterparts). Upon entering this phase, STEP assumes an expanded and more integrated role. It is at this point that data from actual hardware exists to employ the *model-test-model* approach.

M&S is employed in the planning of tests to support a more efficient use of resources. Simulated tests can be run on virtual ranges to predict system performance, assess the system's operational effectiveness, evaluate data collection capabilities and test procedures, conduct rehearsals, and determine if the test limitations (and their effects on the ability to demonstrate whether or not critical system design issues) have been resolved.

M&S is used to provide stimuli with as many operational characteristics as possible or provide a real-time baseline to evaluate the progress of the tests. Test results are evaluated and used for the refinement of the system's requirements and design. Under STEP, they are also used to support the validation of the existing M&S. These models and simulations can then be employed with greater confidence to examine conditions not yet tested.

STEP tools are used to provide data for determining the real component or subsystem’s performance and interaction with other components (e.g., Hardware-In-the-Loop (HWIL), Software-In-the-Loop (SWIL), integration testing), including the operator of the system (e.g., Human-In-The-Loop (HITL)).

Live tests (e.g., open air range tests) are conducted for both DT and OT. Developmental test includes range tests to demonstrate system and subsystem performance and to provide data for the total evaluation effort. This data can be utilized to support the Verification, Validation, and Accreditation (VV&A) process. Operational test is used to assess the operational effectiveness and suitability of a system under realistic operational conditions.

M&S is used during both DT and OT to increase the amount of data and supplement the live test events that are needed to meet their respective test objectives. Through the integrated use of these simulations, a more successful transition is made from DT to OT. Consequently, DT and OT merge into an efficient, effective, and continuous process for evaluation of system performance and operational effectiveness and suitability.

In Conjunction with Recurring Assessments of Military Worth

With the representation of the system at the high level models and simulation, periodically throughout the acquisition process, the current version of the system under development should be reexamined in a synthetic operational context to reassess its military worth. How often this needs to be done is program dependent, but a general guideline would be to conduct such analysis after each significant design change in order to understand the impact of that change on overall theater or battlefield performance and not just on system performance. This is one of the significant aspects of STEP; understanding the answer to the "what difference does this change make" question as the system develops.

Support Throughout the System’s Life-Cycle

During fielding and deployment and throughout the system’s lifetime, the potential for modification of the system continues. Follow-on operational test programs are conducted to assess performance, quality, compatibility, and interoperability, and identify deficiencies in the system. These test programs can use the same set of M&S tools used for earlier Developmental T&E (DT&E) and Operational T&E (OT&E) to assist them in identifying deficiencies and predicting the effects of proposed design changes. New test data can be applied to models to incorporate any system enhancements and to further validate the models.

STEP does not end with the fielding and deployment of a system, but continues to the end of the system's life-cycle. STEP results in a thoroughly tested system with performance and suitability risks identified. A by-product of this process is a set of models and simulations with a known degree of credibility with the potential for reuse in other efforts.

2.3 The Relationship between STEP and the T&E Process

The DoD T&E Process is a five activity iterative process that is repeated many times throughout the acquisition life-cycle. Figure 3 shows a flow chart of one iteration of the T&E process with the five major activities, and inputs and outputs to other research, development, and acquisition activities.

Figure 3. The DoD T&E Process

STEP incorporates the T&E process with the application of M&S tools. In addition to the collection of data to support system evaluation, data are collected to refine the models and simulations using the model-test-model approach. Testing produces models and simulations with increased credibility, and allows for the assessment of system performance in areas and under conditions that might not be otherwise available with conventional testing methods.

Chapter 3 Strategy to Execute STEP

3.l Introduction

In applying STEP within a system acquisition program, one might plan from the macro to the micro level. At the macro level, a strategy to execute STEP would be a useful, and most would say a required, starting point. This strategy might be formulated in the mind of the test planner, or it could be formally developed and then become a part of the TEMP, a test plan, or the M&S plan. The important point is that formulation of the strategy is a natural part of the planning process, and the planner should document the strategy in order to communicate it to peers, superiors, and subordinates in the planning process. The strategy's focus is on the evaluation plan as opposed to a test plan.

The planner should develop the STEP strategy to support and compliment the acquisition strategy. Since the strategy is like a road map, designed to communicate the point of origin, destination, and route, the planner must develop the strategy so he/she can describe it for other planners and the executors, supervisors, and overseers that are a part of the acquisition process. The strategy will serve to let all know what STEP is, where it is going, and the route it proposes. However, the strategy must also be a living entity, changing to keep pace with real world program changes.

3.2 Purpose of the Strategy

The TEMP, detailed test plans, and other test planning documents are already required. . In many cases, M&S plans have also been developed. These documents guide the test process, but the two have often not been synchronized. That is the primary purpose of STEP - to make the test and M&S processes a single one, a process that makes the best use of M&S, live tests, and other tools to collect the data needed to:

- answer the appropriate questions and issues related to system requirements,

- provide that information the decision makers need to make the proper decisions,
- provide program personnel information concerning risk mitigation, and
- verify and validate the models and simulations.

Application of STEP leads to selection of the best available tool to provide that data considering cost, timing, requirements, required information, risk, and resources. That tool may be a test at the component, subsystem, or system level; a simulation; a model; or any combination of test, simulation, and model. However, recent experience indicates that M&S will play a more important, and even a dominant role in providing this data. The decreased resources available for system development has and will continue to result in more extensive use of M&S. The system developer will have to ensure that those models and simulations mature as the system matures. The developer will have to devote resources to mature the models and simulations, and will have to plan for test data to verify and validate the performance of those models and simulations.

It is also important to note that STEP must provide more than data **-** it must provide *information*. STEP provides data concerning system performance, but if failures are noted, the data must be transformed into *information* by assessing the impact of that failure and providing information on its relevance to system performance, developmental risk, the acquisition strategy, and at a minimum, proximate cause.

As STEP progresses, the models and simulations are matured resulting in less error noted between predicted and test data. Consequently, as illustrated in Figure 4, the confidence bounds on the information which is developed become smaller, giving the decision maker increased confidence in his decisions.

Figure 4. Bounding Confidence on Information

3.3 Developing the Strategy

In planning for STEP, the planner should tailor the strategy to fit the:

- type of system (missiles, ships, and satellites require different strategies),
- decision maker’s information requirements (what is required and when),
- verification and validation data requirements,
- risk (which subsystems or technologies present the greatest risk to technical and operational success),

- resources (providing value added within budget constraints),
- timing, and
- tools available.

Type of system

The type of system being developed will influence the acquisition strategy, funding profiles, schedules, and all other development factors, to include the STEP strategy. Additionally, the availability of M&S tools is related to the type of system. For example, developers of missile and satellite systems have long made extensive use of M&S. Therefore, many of the tools such developers will need for the future already exist. However, other communities have not made such extensive use of M&S and may have to develop many of the M&S tools needed to support STEP.

Information Requirements

The information required to support the acquisition decision process will influence the strategy. At the higher levels, the information requirements will be directed at answering issues on cost, schedule, performance, technical risk, and program risk. However, more detailed information is required to address these issues in terms of Measures of Effectiveness (MOEs), Measures of Outcome (MOOs), and Measures of Performance (MOPs). A good understanding of the information requirements can be obtained by reviewing the Critical Operational Issues and Criteria (COIC), Critical Technical Parameters (CTP), and system specifications. Once the general need is determined, existing models and simulations that are capable of addressing these questions must be identified. If the existing models and simulations cannot address the questions, then modifying existing tools or developing new ones must be considered.

Data for V&V Requirements

The STEP strategy must provide for models and simulations to mature as the system matures. The strategy must verify and validate the models and simulations in order to measure that maturity and ensure that they do what the model and simulation designer intended and that they accurately represent the real world for the intended purpose. To accomplish this in a fashion that supports the STEP, planning is essential. This task requires test data obtained from actual hardware to verify and validate the models and simulations. It also requires qualified people who are given the time and data necessary to accomplish their task. The strategy should consider who should accomplish the verification and validation (V&V). Should it be accomplished by Government personnel, or should it be done by contractor personnel? If the contractor personnel work for the system developer, should the strategy provide for oversight by an independent party? Recognize that persons familiar with the models and simulations will complete the V&V more efficiently than persons who are unfamiliar with them, but there are risks involved in that approach for which the planner must account. However the V&V is done, the goal is to make the model more accurately represent the actual system and operating environment as time progresses.

Risk

Technical risk in the system development is a major issue in development of the strategy. The M&S effort may need to focus on those areas of the system presenting the greatest technical risk. For example, in development of a missile with new seeker technology, the planner may want to provide for a HWIL facility as early as possible. This would provide a method to test the seeker technology using actual hardware and software with simulated environment and threats. Other STEP tools may be able to accomplish similar goals, but the strategy must cause the planners to provide the facilities, people, and resources to accomplish the goal. M&S by itself will not identify unknown-unknowns. Other STEP tools, to include field testing, are needed to accomplish this.

Resources

The STEP strategy must add value for the resources available. The resource requirements needed to accomplish the STEP goals must be communicated to the resource planners. The resource planners need this information early in general form and in more detailed form as time progresses. STEP strategy development involves trade-offs among the STEP tools to accomplish STEP objectives. Since resources are constrained, the most efficient tools must be selected to provide the quality information required by the decision maker.

Timing

The program schedule must be considered in developing the STEP strategy in order to provide the data the engineers and analysts will need at the time they will need it. Timing is critical to development of the strategy. When is the information required, and how much time is available to find, develop, mature, and use the M&S tools needed? The information is not relevant if it comes after the decision is made. The STEP strategy must be designed to provide the best information reasonably available when the decision is to be made.

Tools

M&S tools range from live tests (themselves a simulation) to pure simulations. This includes HWIL, SWIL, HITL, SIL, and ADS. These tools can provide data from the engineering, engagement, and campaign levels. The planner must consider the types of data needed to provide the requisite information, and develop the strategy which will provide that data. Chapter 5 will discuss these tools in more detail, to include the factors to consider in planning for their use.

Chapter 4 STEP and the TEMP

4.1 Overview

STEP will have a significant impact on the methods by which traditional test programs are conducted. The impact will be reflected in test activities such as evaluation planning, test planning, and documenting that planning in the TEMP. The TEMP is a "living" document which must accommodate ongoing changes to the acquisition program. The TEMP documents the integration of M&S with the T&E process through STEP and is iterated in conjunction with the development of the new system.

STEP begins with the examination of mission need and the development of operational requirements. M&S tools are used to examine anticipated mission scenarios, threat environments, analysis of alternatives, and technical requirements to help provide a clear understanding of the system capabilities needed. Focused understanding of the mission need, employment context, and requirements aids in the development of precise, testable, technical parameters, operational issues and criteria, and achievable thresholds and objectives. This solid foundation leads to the development of clearly defined DT and OT programs as well as realistic schedule, data, and resource requirements.

The following sections describe how STEP is integrated into the parts of the TEMP.

4.2 System Introduction

This part of the TEMP documents the measures of effectiveness and suitability derived from system requirements. The traceability of test results to requirements is key to the successful accomplishment of the goal of verification of system requirements. By utilizing the iterative method of STEP, the traceability of requirements from the MNS, ORD, and performance-based specifications to the MOEs and MOPs is clearly discernible. The comparison of simulation output to the MOOs, MOEs, and MOPs can bring early insight to the program not previously available with "traditional" test methods requiring physical system prototypes. The program can use these comparisons to examine the test methods, expand the envelope of the test environment, and reexamine the ability of the system to meet the requirements. The test planning should give careful consideration to the selection and collection of data for these comparisons

4.3 Integrated Test Program Summary

The integrated test program schedule is the union of all the individual test program parts. It reflects the synergy of the DT, OT, live-fire tests, requirements, resources, time, and funding that culminate in an achievable test program. In order to do this the Program Manager must have a clear understanding of the method by which the system will be developed, tested, and evaluated. STEP integrates test and simulation, and this must be reflected on the integrated test schedule in order to facilitate a successful evaluation strategy.

STEP impacts the integrated test schedule by applying M&S early and throughout the development and test process. By using simulation to accomplish many of the required tasks (reducing the need for expensive "live" assets), the program may be able to avoid expenditures in resources and time. However, the Program Manager must take into account the spectrum of combinations of M&S and test available to the system under development and assess the applicability of those resources. Development of new M&S resources, modification of existing M&S tools, VV&A of appropriate models and simulations, and the availability of data for these resources will all have a significant impact on the development of a realistic, achievable evaluation strategy and schedule.

Therefore, it is imperative to the success of the program that any development of M&S tools be included on the integrated schedule. Design and coding reviews, planned incremental releases, site installation, and related VV&A activities must also be included. Likewise, modifications and VV&A of existing M&S tools as well as data collection need to be delineated on the schedule to ensure adequate time and resources to accomplish all the tasks required.

4.4 Developmental Test and Evaluation Outline

The majority of the iterative work of STEP is accomplished during the DT program. It is here that STEP uses M&S to incrementally develop and test the system, making refinements to the system as well as the M&S tools.

During DT data is collected for system performance verification and refinement of requirements, and models. Careful planning must be done to ensure that appropriate tools are selected to conduct the tests and that the right data is collected to fully assess the achievement of the objectives. The TEMP should reflect a logical progression of testing which incrementally develops the system and the models as well as provides data for validation of system requirements.

Initial iterative STEP phases which refine the requirements and objectives of the system provide an examination of existing M&S and its appropriate application to the evaluation strategy and the test program. Use of M&S during the early phases of system development provide information on risk areas that need to be addressed, "data holes" where M&S tools may need to be developed or enhanced, and areas where live testing can be augmented by simulation. This information leads to the development of test methods and data sources/types and the correlation of associated system requirements and objectives.

The validation of M&S tools is conducted throughout the DT program. Data collected from simulation is compared with all available test data. These comparisons allow for refinement of M&S, tests, and data collection practices which improve both the system development and system performance evaluation. In addition this data comparison provides the basis for the validation of the M&S tools. The TEMP should reflect the method with which the VV&A is to be conducted. Much of the information gathered during the initial phases of STEP can provide the required justification for the credible use of M&S.

The STEP method allows for the test program to "expand the envelope" to examine operational issues early in development. Through the use of M&S the program can examine system performance from previously unavailable vantage points. M&S allows for the controlled introduction of unique variables without increasing overall program risk. These initial operational insights can provide invaluable information on the data collection requirements, performance requirements, thresholds and objectives as well as the applicability of MOEs and MOPs. The TEMP should reflect the development and execution of these expanded test events.

4.5 Operational Test and Evaluation Outline

The application of STEP facilitates the early operational assessment of system capabilities. These early operational assessments coupled with the results from the iterative STEP phases of DT provide data to support the certification of readiness for the system to proceed to OT.

A key issue with STEP is the extent to which "live" operational testing can be augmented with simulation. During the conduct of DT, the M&S tools have gained credibility. The up-front use of M&S aids in the assessment of the operational impact of the system under development, human factors interactions, and OT requirements. The use of accredited models and simulations to assess the capabilities of a system greatly reduces the risk associated with OT scenarios. As with DT, careful planning must be done to ensure that the appropriate data will be available from all aspects of STEP to assess the satisfaction of system requirements and the COICs. The method for obtaining VV&A of the key M&S tools must be stated.

STEP expands the arena of operational tests that can be conducted with the use of M&S. Engagement- and battlefield-level missions can be conducted early on to develop interim operational procedures and training requirements. STEP also allows for interactive freeplay in the operational scenarios which increases the operational as well as the human factors/training information gained from the test program. This flexibility is gained from the iterative nature of STEP and the careful gleaning of information during the previous phases to validate and accredit the M&S, verify system performance, and refine/validate the system requirements.

It has been stated that anything short of war is a simulation. The extensive use of an integrated set of M&S tools and the iterative STEP process provides an environment within which the most extensive data collection, both developmental and operational, can occur while reducing overall program risk and increasing system effectiveness.

4.6 Test and Evaluation Resource Summary

The following STEP-related resources need to be identified in this part:

- **Test ranges/facilities including Major Range and Test Facilities Base (MRTFB), industrial and academic facilities**
- **Threat representations including simulations**
- **M&S resources to include VV&A**
- **Unique instrumentation and targets.**

STEP changes the focus of the types of resources required for test programs. The use of up-front and continual simulation may reduce the need for "live" test assets and focus the need for test assets to the key critical tests. The use of simulation can also increase the efficiency of test programs by allowing for "dry runs" which can identify procedural and data collection difficulties in advance. Interactive M&S tools can aid in greater understanding of the interaction of systems with operator capability.

The M&S tools used in each stage of the acquisition life cycle may differ as the program matures. However, the resources should not be abandoned, but archived into a repository to be used at later stages in the life cycle of the program or utilized by other programs. The listing of resources should also reflect the estimated resources required to VV&A models and simulations as well as the resources required to obtain and maintain the M&S tools and associated data.

Data collection and validation are critical to the successful ongoing use of models. A genealogy of the M&S development and VV&A depends on the data collected to provide the substantiation for each stage of the VV&A. The data collected must also be archived for future use by the current system or by other systems desiring to utilize existing M&S technology.

While STEP may reduce the number and extent of live resources required for testing, care should be taken to ensure that adequate resources are available for the life of the program. The iterative nature of STEP provides the means of ensuring that the appropriate and testable requirements are developed. It can also aid

in the development of a complete resources plan to ensure a complete and accurate test program by allowing for dry runs to determine the overall resources need.

Chapter 5 STEP Tools, Standards, and Resources

5.1 Introduction

Many of the tools needed to support STEP have been used to support the systems acquisition process, both in DoD and industry, for a long time. As dramatic advances in the supporting technologies made those tools more powerful and less expensive, and as declining resources and changing priorities made it essential to find better ways to develop and field new systems, the use of these tools and of associated improved processes that exploit their contribution has expanded rapidly throughout the system acquisition process.

The sections that follow outline three areas of interest in the application of tools supporting STEP, namely Tools, Standards, and Resources. These tools are one vital element of STEP, but are not the universal answer to all T&E problems. STEP tools do provide alternative sources of assessment data and a degree of repeatability, time-stepping and analysis not available in live tests. STEP seeks the total integration of productive M&S tools to optimize T&E in support of risk mitigation in the acquisition process.

5.2 Tools

The description of STEP tools that follows is not all-inclusive, but is provided to show the types that exist and illustrate potential applications to STEP.

<u>Models and Simulations:</u> The development of the models that

describe the system and the M&S that predict its performance is driven by the identification and refinement of the system requirements. For example, the Mission Needs Statement (MNS) for a major weapon system initiates the selection or development of models and simulations that reflect the performance of the proposed system in theater/campaign, mission/battle, and engagement-level conflicts. The top-level system requirements of MAIS drive the development of top-level models of computer hardware and software, data, and telecommunication systems that can be used in simulations to predict functions such as collecting, processing, transmitting, and displaying information.

M&S can be used to design a better test program, add realism to test scenarios, extrapolate results of testing, and explain aspects of system performance observed during testing. This can reduce time, resources, and risk to an acquisition program. In some cases, M&S is the only way to conduct system assessment and is the only way to generate "reproducible" scenarios and conditions. The major limitation associated with models and simulations is credibility which is addressed in more detail in Chapter 7 of this document.

<u>ADS</u>: Advanced Distributed Simulation. ADS is an environment in which simulations are linked to produce large synthetic environments within which large numbers of subjects can interact in real time. The principal characteristics of ADS are that participating simulations are physically separated, are linked electronically, and share a common view of their electronic environment. The responses elicited from each simulation are seen, interpreted, and acted upon by the other simulations in near real time.

ADS offers the potential to link multiple non-collocated T&E for an evaluation. This could include multiple systems being linked and operating within a realistic synthetic environment, thus creating a virtual system. The benefits of this linkage are early interoperability and compatibility evaluation, and early user input to the system operating in concert with other new systems in a synthetic environment. Cost savings could be realized through early testing, reduced need for transportation and repositioning, and reduced acquisition cycle time. Limitations of ADS include latency, or the communication delay between participants, and the immaturity of the technology. The challenges of VV&A are compounded by the distributed nature of the ADS network; the linked network will require VV&A in addition to each stand-alone M&S participant.

CFD: Computational Fluid Dynamics. CFD is a numerical approach for modeling the dynamics of a fluid flow in and around solid objects. In aircraft applications, it can be applied to model stores separation to analyze aircraft-store loading, safe carriage and separation, safe escape and ballistic accuracy. The cost associated with certification of stores for release from an aircraft is very high. CFD is often used as a tool in direct support of wind tunnel and flight testing. Integrating the modeling tools directly with ground and flight tests enables the tester to design a better test program, validate and/or extrapolate the results, and assist in decision making for a more efficient or effective test. CFD can be applied to investigate anomalies observed during wind tunnel testing at significantly reduced cost over repeat wind tunnel trials. The principal limitation of CFD is credibility. Optimal application of CFD is in conjunction with wind tunnel and flight test/analysis effort employing the model-test-model approach.

Simulators: Simulator has a number of meanings. It is a family of equipment used to represent threat weapon systems in testing and training. A threat simulator has one or more characteristics which, when detected by human senses or man-made sensors, provide the appearance of an actual threat weapon system within a known degree of realism. It is also a human-in-the-loop device that provides the conditions and environment of a system to accurately produce aspects of the systems performance and operation to conduct training and develop tactics.

Threat simulators are useful to testers as a means to mitigate limitations to the scope of testing because actual threat systems operated by trained "enemy" personnel may not be available for realistic testing. Threat simulators can be emissions, signatures or radar returns synthetically injected into a controlled laboratory environment, or they could be friendly units playing the role of threat forces in doctrine or performance. In some cases, threat hardware is used. M&S "simulators" may produce their greatest utility in providing the electronic means to generate sufficient numbers of threat forces to provide a meaningful evaluation of new system capability.

Simulators in the second meaning is are commonly employed as training devices, but have numerous applications in support of testing, system operation, and tactics. Simulators developed concurrently with the system provide the means to obtain user feedback on operational and human factors issues. When integrated with other systems via ADS, great insight into the operational aspects of the system can be obtained.

Stimulators: A stimulator is a simulation used to provide an external stimulus to a system or subsystem. The output of the simulation is used to "stimulate" the system (hardware and/or software) being evaluated for purposes of analysis. The unit under test may be in a HWIL configuration, test stand, or live field conditions. Limitations include the degree of realism of the stimulator, lack of tactics applied, and total threat system representation in fit, form and function to replicate the threat.

HWIL/SWIL: Hardware/Software In the Loop. HWIL/SWIL is a hybrid simulation that includes actual system (prototype or production) hardware or software in conjunction with digital models and external stimuli to demonstrate the operations and functions of the hardware/software within an environment simulating actual operating conditions. Functioning of the simulation with the hardware or software is often accomplished in an integration laboratory.

HWIL/SWIL can be used to demonstrate new technology; evaluate designs, concepts, and prototypes; and show the integration of hardware and software. It allows early evaluation without the expenditure of live test resources, facilitates live test development by identifying desired test conditions, and can be used for development of data collection plans. The lab environment may provide for easier data collection as a result of better access to components.

In the case of software testing, SWIL may provide the only method to examine complex software and the algorithms and logic flow of programming in adequate detail. If the operational conditions or environment of live testing is such that conditions necessary to manifest a fault do not occur, the fault may not be detected. If undetected, the fault can be duplicated or promulgated across other interfacing systems. Correction of faults after software release and integration in deployed systems is difficult and expensive. SWIL is the best opportunity for fault detection and software risk reduction prior to release, and is one of the most cost effective methods of fault correction.

For early phases of operational testing, HWIL can demonstrate potential effectiveness and suitability for designs that exist in a pre-production or breadboard stage of development. User feedback on HWIL/SWIL performance can be used to improve the design early in the process. The reduction in risk by the early system integration can be of significant value to the Program Manager. In some programs, complete HWIL simulations housed in elaborate System Integration Laboratories have provided fully integrated system testing prior to platform installation and offer the additional benefit of crew performance evaluation. The principal limitation to HWIL/SWIL simulation is realism in the operating environment and credibility of the system representation. Use of ADS to bring realistic synthetic environments into the testing loop and robust VV&A of the system representation can mitigate these problems.

SIL: System/Software Integration Laboratory. A SIL is a facility that supports the integration of system components and/or software in a laboratory environment for development, experiments, and testing. The integration laboratory "simulates" (or replicates) a system to a known extent and allows the modification/addition of component hardware/software for use without many of the restrictions or difficulties that would be encountered using actual system hardware or host platforms. Testers will obtain the benefits of HWIL/SWIL testing. The SIL is the physical support structure and components that make

HWIL/SWIL testing and evaluation possible.

ISTF: Installed System Test Facilities. ISTF are facilities where entire systems or sub-systems get their first workout in the environment in which they will operate (e.g., inside an aircraft). A full capability ISTF has the ability to mix a complete spectrum of players from synthetic (digital models) to real (actual hardware) to hybrid (a combination of both); the ability to provide multi-level threat simulations (open-loop and closed-loop signal simulators, including actual or simulated threat system hardware); and the ability to provide simulations of all C3 elements a system would be expected to operate in the real world. The Navy’s Air Combat Environment Test and Evaluation Facility (ACETEF) at Patuxent River, MD is an example of an ISTF.

Measurement Facilities. Measurement facilities are used to quantify or measure parameters (such as thrust, radar cross section, and drag) of a test article in precise terms. Examples of such facilities are wind tunnels, radar cross section facilities, antennae pattern ranges, and engine thrust stands.

Wind Tunnel: A chamber through which air is forced at controlled speeds so that its effect on an object can be studied. A wind tunnel can be utilized at different times in a system’s development to analyze air stream effects. This testing is especially relevant for flight safety evaluation and issue resolution. This evaluation can be augmented with CFD, flight testing, and captive carry evaluation. Limitation of wind tunnel use are the cost and availability of the facility and limitations on weapons system size that can be accommodated. Due to the cost of development and operation, most wind tunnels cannot accommodate full scale systems. This forces scaling down of systems to less than full scale models for testing and imposes limitations on conclusions drawn from wind tunnel tests. The use of wind tunnels is a proven evaluation tool in the development of aerodynamic bodies prior to actual development for flight testing.

Anechoic Chamber: A facility that provides an essentially echo-free environment at various electromagnetic frequencies for laboratory measurements including radar cross-section measurements, antenna pattern measurements, and passive radar augmentation measurements. Anechoic chambers provide a "pure" environment for testing, and the additional benefit of secure testing with no external emissions. Weapons system size limitations may preclude some types of testing.

Captive Carry: Captive carry of stores or vehicles is a non-destructive means of obtaining data on air-capable vehicles. Instead of free-flying the vehicle, it is mounted on a parent aircraft. This "host" aircraft provides the motion, power, and support services required to operate the "guest." Additionally, the host aircraft often collects the data obtained for later analysis. A captive carry store or vehicle is theoretically not damaged by the captive carry evolution and can be examined, analyzed, evaluated, and reused.

5.3 Standards

HLA: High Level Architecture. The HLA has been set as the DoD standard architecture for M&S throughout DoD. HLA provides a common architecture for reuse of simulations. It is based on the premise that no single model or simulation can satisfy all uses and users in DoD at all levels of resolution. An individual simulation or set of simulations developed for one purpose can be applied to another application under the HLA concept of the federation: a composable set of interacting simulations. The intent of the HLA is to provide a structure which will support reuse of capabilities available in different simulations, ultimately reducing the cost and time required to create a synthetic operating environment for a new purpose.

JMASS: Joint Modeling And Simulation System. JMASS is a program to develop and deliver a distributed, object-oriented M&S architecture and system focused on the tactical level of war (mission and engagement level simulation). It is designed to provide a flexible, standardized M&S tool to support a wide range of simulation requirements throughout the life cycle of a weapon system. It supports concept design, trade studies, OT&E, DT&E, and tactics analysis and development. The JMASS software is aligned to the following five M&S activities:

- Develop mode is where individual components are created using the JMASS visual tools. These objects map to the real world and support object based simulation. JMASS tools automatically produce code compliant with the JMASS software structural model.
- Assemble mode provides the tools to create JMASS "players" by combining individual objects into more complex objects. Model developers can access the JMASS Modeling and Simulation Reuse Library for components and/or develop new components to build the model which meets the specific analysis requirements.
- Configure mode supports JMASS users as they create a

specific simulation scenario, lay down the individual components in some physical space, and array data against the variables in the simulation.

- Execute mode allows the simulation to execute as it was configured. Experiment management tools allow multiple runs to be made with different variable settings.
- Post Processing mode supports the analyst's understanding of the simulation results. Data can be presented in a variety of tabular and graphical formats.

5.4 Resources

MSOSA: M&S Operational Support Activity. MSOSA is a service that supplies M&S resource information to DoD customers. It catalogues M&S resources, databases, exercises, capabilities, and points of contact and makes this information available to the DoD community. MSOSA resources link to each of the Services, the Information Analysis Centers, and other specific M&S sources. Through an on-line database, information is readily available. MSOSA also provides M&S expertise and direct access to the DoD M&S "corporate knowledge base."

Program Managers utilizing MSOSA may be able to benefit from the reuse of existing M&S products and databases; obtain better M&S exercise support; obtain easy access to information on M&S policies, activities, and initiatives; coordinate with on-going M&S "events"; and experience cost savings through asset sharing. MSOSA builds their databases based upon the needs of the community, and expands their capability based on contacts made and requests for services. MSOSA provides a staffed help desk and 24 hour on-line service via the Internet. MSOSA can be contacted on e-mail at: msosahelps@msis.dmso.mil.

MSRR: Modeling and Simulation Resource Repository. A linked system of resources (servers, classified as nodes) that allow connectivity, re-use, and sharing of M&S resources to support communication and information sharing. A prototype MSRR, leveraging the existing technologies of inter-networking is established to provide a number of useful services, including software tools, electronic documents, Government off-the-shelf applications, a repository of M&S data models, and directories and catalogs.

MSEA: M&S Executive Agents. MSEA are DoD Components who have been assigned management responsibility by USD(A&T) for the development and maintenance of a specific area of M&S application, including relevant standards and data bases, used by or common to many models and simulations. The current MSEA are: Air and Space, under the Air Force at AFCCC; Oceans, under the Navy at OP-096; Terrain, under the Defense Mapping Agency at the Terrain Modeling Project Office; and Threat/Intelligence, under Defense Intelligence Agency (DIA).

Chapter 6 STEP Application

6.1 Introduction

In Chapter 2, STEP was defined and the STEP concept and process was presented with the perspective of developing an interdependent set of credible M&S tools that could be used to support the T&E of acquisition systems. The primary product of STEP was identified as information for acquisition decisions. Beneficial byproducts of STEP are models and simulations of known credibility available for subsequent use on other programs.

Figure 5 represents the flow of STEP across the acquisition activities (right face of the cube), the M&S levels (front face), and the STEP tools (top face), to support the decisions facing the decision makers as part of the acquisition activities (right face). This chapter discusses the application of STEP as it implements this relationship across the acquisition phases. The portions of the cube in Figure 5 which apply to each acquisition phase will be highlighted and discussed in the following sections.

Figure 5. The Application of STEP

6.2 During Concept Exploration

During Concept Exploration (CE), Analysis of Alternatives (AOAs) are developed to aid and document decision making by illuminating the relative advantages and disadvantages of the options being considered. Analysis performed shows the sensitivity of each alternative to possible changes in key assumptions (e.g., threat) or variables (e.g., selected performance capabilities). In addition, the most promising system concepts are defined in terms of initial, broad objectives for cost, schedule, performance, software requirements, trade-off opportunities, overall acquisition strategy, and T&E strategy.

During CE, through the application of STEP, theater/campaign- and mission/battle-level models are used in constructive and virtual simulations to explore variations in system performance. These simulations can be used to generate data for trade studies, requirements evolution, development of performance objectives, and development of the evaluation strategy and the test strategy for the alternative concepts. See Figure 6**.**

Figure 6. The Application of STEP During Concept Exploration

During this Phase, M&S will be used extensively to assist in developing information required by acquisition decision makers and refining requirements, since most programs will not have a physical system to examine at this time. The Program Managers can use M&S to help them understand and refine the requirements, and develop testable measures. The system is represented in early simulations by assumed values for system performance, which are based on mission need and the system-level requirements as documented in the MNS and the Operational Requirements Document (ORD). The consequences of variations in system performance can be thoroughly explored, by conducting simulations that follow established scenarios and that are programmed to output predicted values for MOOs, MOEs, etc. Simulation output can be used in support of trade-off studies, to determine if requirements are testable, and in establishing system performance and operational test criteria.

During CE the T&E community will begin developing the M&S and T&E plans to include establishing the strategy, identifying major test objectives, simulation and test events, resources, and timing. This includes determining how the test objectives will be met. It includes conducting a search for existing theater/campaign-, mission/battle-, engagement- and engineering-level models and simulations, evaluation tools and test facility resources (e.g., HWIL, integration laboratory, flying testbed, open air or range) needed to execute the tests, conduct the evaluations, and support the simulations, and to plan for their acquisition or development. It also includes a plan for the logical progression of test events to gather the test data appropriate to the degree of system development.

The plan developed will attempt to leverage the needs of the T&E community with those of other acquisition activities (e.g., design, risk management, trade-off studies) and the training community. This plan, when implemented, should strive to provide the program with a common tool set, providing consistent results and potentially reducing the cost of the program.

6.3 During Program Definition and Risk Reduction

During the Program Definition and Risk Reduction Phase, theater/campaign-, mission/battle-, engagement-, and engineering-level models and simulations are used in prototyping demonstrations. M&S conducted during this phase produce information for risk reduction and for refinements of the assessments of the alternative concepts. See Figure 7.

The most detailed descriptions of the system are embodied in the myriad of Engineering-level models and simulations that describe the system, its subsystem and components, and their interactions with the environment, the war fighter, and the intended threat. Engineering-level models and simulations are concerned with phenomenology (e.g., aerodynamics, acoustics, electromagnetic wave propagation, fluid flow, hydrodynamics), physical design, performance, cost, manufacturing, and supportability of the system. During this phase, they are used in predictive simulations to evaluate the system's compliance to design specifications and the system's performance (e.g., predicted values for MOPs and Technical Performance Parameters (TPPs)).

Figure 7. The Application of STEP During Program Definition & Risk Reduction

Engineering-level M&S can support virtual prototyping of the system producing predicted performance values long before real prototypes are built, and minimize the number of real prototypes that must be built to effectively evaluate the system.

Predictions from engineering-level M&S are also used as the basis for the system's performance as represented in engagement-level M&S (i.e., the predictions replace the assumed values). Engagement-level M&S provide information on the effectiveness of the system under test against a specific target or enemy threat, thus providing MOEs at the system-on-system level.

EOAs in support of Milestone II have traditionally been performed on the physical prototypes. STEP ensures the PM has an integrated set of models that can be used in simulated test scenarios so EOAs can be made without the need to build physical prototypes. These EOAs can even be performed during CE using virtual prototypes to support Milestone I decisions.

6.4 During Engineering and Manufacturing Development

During Phase II, Engineering and Manufacturing Development (EMD), the most promising design approach is translated into a stable, interoperable, producible, supportable, and cost effective design. During this phase, the T&E community demonstrates system capabilities through tests and evaluations. Models and simulations at all levels in the hierarchy are used in constructive, virtual, and live M&S and tests to verify the system's design, verify that design risks have been controlled, certify readiness for operational testing, and determine if the system is operationally effective and suitable for the intended purpose. See Figure 8.

Figure 8. The Application of STEP During Engineering and Manufacturing Development

Initially, the test managers may utilize a number of digital M&S tools to provide guideposts to testing and predictive results. (These same M&S resources were used to provide the guidance for the development of the test program and, in the early phases, provide the structure for individual test program planning.) As the test program progresses, the use of STEP tools expands to include wind tunnels, stimulators, digital mock ups, and HWIL testing, and then system integration and installed system testing. Finally, M&S and tests using scale models, exercises, open air ranges, and distributed interactive simulations may follow.

Validated input data for simulation purposes is key to the use of accurate and appropriate M&S tools in the test program. Data repositories are still in the infancy stages for most M&S tools. However, validated data for environmental factors, human interaction and some system performance do exist. In any case, it is important to be able to trace the source of input data for M&S tools as questionable data can introduce unnecessary risk to the program.

The method of data collection is often dictated by the instrumentation and recording equipment available at a given facility. Planning for data collection includes researching the existing data collection methods and opportunities for development. M&S can aid in planning individual tests by identifying inconsistencies in data collection, "bottle necks" in the data stream, and instrumentation problems. In addition, M&S can be used to explore and/or develop new data collection technologies and methods.

Methods for data validation are an emerging technology area. The method of validating M&S and test output data must be clearly defined and consistent with the method of test. Actual test data should be captured and used, where appropriate, to validate the outcomes of M&S. The data collected from both actual and "simulated tests" should be compared to expected outcomes produced in earlier activities. Comparison of these data help to indicate whether the program’s performance and test methods are on track.

An important element in the STEP philosophy is the iterative efforts to improve the models and simulations with representative system test data. This allows progressive improvement in the M&S which supports the tests with performance characteristic predictions used in planning the continuing tests and in test risk assessment. Test results which depart significantly from the predictions should result in a suspension of testing until the M&S can be updated and new predictions generated. These new predictions are then used to guide the selection of the most useful test points when testing resumes. The final result when testing is completed and the final M&S updates are made are validated M&S resources which will sustain the weapon system throughout its life cycle.

Pre-test planning includes the use of M&S to exercise the test, data collection, data reduction, synthesis, and evaluation procedures. By simulating this plan for the test, data collection, and data reduction procedures, the T&E planner can assess the capability of the data collection and processing method to produce the appropriate output for evaluation. The data can be reviewed for holes, determining which requirements may have been untested and the best method of obtaining that data. This review may also reveal data requirements which cannot be collected during actual tests and require additional simulation or, conversely, reveal data requirements which should not be collected through simulations and for which actual tests must be conducted. This evaluation also may reveal areas where the actual test data and simulation data contradict each other and additional investigation is required to establish the source of the contradiction. The evaluation process helps to refine the M&S and test program or the data collection method. It helps to determine what data should be provided for the evaluation process and in what format, and may also reveal additional sources of data required. M&S resources are key in the evaluation process, and the use of M&S results and test data in support of the evaluation is a significant task.

Finally, it is important for the Program Manager to plan for the feedback of simulation and test data back into the models and simulations. The feedback from tests to the simulations are not only required for VV&A, but can provide insight into areas where M&S and tests may be lacking - gaps in requirements testing, refinement of environmental constraints, and adjustments to modeling logic or algorithms.

The test results are evaluated and used for the refinement of the system's requirements and design. These results are also used to validate and support accreditation of the simulations for additional use. With the improved models, engineering-level models and simulations are run outside the test conditions to obtain predictions of performance over a wide range of conditions not covered by test (due to a variety of programmatic issues, such as limited funding, political and environmental issues, and safety constraints). These predicted results constitute improved system performance parameters for use in upper-level models and simulations. For example, they are used in engagement-level models and simulations to improve the fidelity of the values for MOE predictions at the system-on-system level.

Concurrent with DT, OT&E evaluators prepare OAs of the system's performance and suitability. These OAs are made on prototypes and low-rate initial production (LRIP) units of critical systems, subsystems, and components to ascertain if the risks and operating environments are within acceptable limits. OAs are also performed to determine if test limitations, and their effects on the ability to demonstrate whether or not critical system design issues and design risk, have been resolved. M&S can be used as a predictive tool to extend the range of the assessments made on physical prototypes and LRIP units.

Operational testing provides data on the actual hardware and software in an operational environment. Operational tests are conducted to determine the operational effectiveness and suitability of a system under operationally realistic conditions.

Production units (LRIP units initially) can be embedded in the established engineering-level and engagement-level M&S as appropriate to the test scenarios and linked through ADS as needed to execute their respective parts in the established test scenarios. Live tests are run and the results compared to the corresponding predictions from constructive M&S for model and simulation maturation and as input for the continuous VV&A process. Models and simulations are run to augment and extend the test results, and to provide more accurate performance parameters for use in higher-level M&S. Information supports decisions to proceed beyond LRIP.

Carefully planned operational tests can provide data at the mission/battle-level. For the most part, evaluations of operational effectiveness and suitability (MOEs and MOSs), and conflict outcomes at the theater/campaign-level (MOOs) can only be provided by M&S.

6.5 During Production, Fielding/Deployment and Operational Support

During Production, Fielding/Deployment and Operational Support, the entire accredited, interdependent set of STEP tools is available for use as appropriate. See Figure 9.

Follow-on operational test programs are conducted to assess performance, quality, compatibility, and interoperability, and identify deficiencies in the system. These follow-on test programs will use the same set of STEP tools used in earlier DT&E and OT&E. Furthermore, they will conduct tests to collect data for updating/refining the models, and as input for the continuous VV&A process, to ensure that the latest and most accurate M&S are available for use in future evaluations and to support other acquisition activities and the training community.

Data representing the system can be converted for use by numerically controlled machinery, quality control, or other processes during manufacture. Changes introduced during production must also be reflected as changes to models and data sets to ensure that fully representative tools and documentation are available. This process of updating models and data sets is critical in

Figure 9. The Application of STEP During Production, Fielding/Deployment & Op Support

resolving design problems found in service and in making modifications to the system throughout its life cycle. Interoperability of the weapon with other systems can also be verified using the M&S assets. The resources can then be transitioned to support new systems under development.

The M&S verified and validated during development and testing have wide use during fielding and deployment. Training simulators for system operators and maintainers must produce as realistic an environment and reproduction of system functions as possible. This will most likely be based upon or made up of subsets of the earlier M&S resources. Operating specifications and characteristics presented in manuals and electronic mission planning aids are also derived from such M&S assets. Models of system durability and longevity will be used in determining timing of servicing, planning for parts availability, and other logistical support requirements. Records of system usage, failures and maintenance actions are folded into certain models of system performance and longevity. This allows necessary adjustments to operating procedures and limitations, plus periodic depot maintenance timing and planned work, to meet actual in-service behavior. Data sets or M&S are also rolled up as increasingly simplified, but authentic, representations of the system in wargame exercises from the mission to the campaign level.

Chapter 7 M&S Credibility

7.1 Introduction

The success of STEP depends greatly upon the credibility of the STEP tools which are employed throughout the process. The credibility of the M&S and supporting data is measured by a structured validation and verification (V&V) process. The M&S V&V is approved as acceptable for use in a particular application by *accreditation*, with the entire process known as VV&A. Similarly, data used in the M&S is *certified* through a process called VV&C. Ensuring that M&S and data used in or generated by STEP has been properly accredited and certified must become a rigid discipline. This chapter sketches VV&A and VV&C, the requirement for careful software configuration management, and some of the practical issues involved in the use of M&S tools.

7.2 M&S Accreditation

Verification focuses on M&S capability while validation focuses on M&S credibility. Verification is the process of determining that a model implementation accurately represents the developer's conceptual description and specifications. Validation is the process of determining the degree to which a model is a sufficiently accurate representation of the real-world (the subject system and the operating environment) from the perspective of the intended uses of the model.

Accreditation reflects a decision to use an M&S tool for a specific purpose or application. The decision is supported by the V&V and certain documentation. The process that leads up to an accreditation decision gathers all the information about specific model or simulation capabilities relative to the requirements of a specific application. This information includes the V&V results, but also includes such things as simulation run time, number of simulation operators required, the simulation's history of use, documentation status, configuration management, and other factors. Documentation includes the *V&V Plan*, the *Accreditation Plan*, associated reports, and more.

All M&S are driven by data, either as direct inputs by the user or as embedded constants that drive simulation characteristics. Both the data producer and the data user are involved in data verification, ensuring that data meets specified constraints defined by data standards, and that the data are transformed and formatted properly. Likewise, both are involved in data validation, assessing whether the data is appropriate for use in the intended model within stated criteria and assumptions. There are data V&V processes and procedures which parallel those for M&S V&V described previously. For data, the decision to use data for a specific application lies behind the certification. The configuration management, documentation and similar issues relevant to M&S V&V also underlie data VV&C.

Software configuration management (CM) is of critical importance to STEP. Configuration management is a development life cycle process through which the integrity and continuity of software development, upgrades and maintenance are recorded, communicated, and controlled. If there is no effective CM, a user cannot be assured what version of the M&S an application is using or what code, hardware and/or data is being used. Good CM usually implies good documentation. Poor or no CM leaves any M&S documentation suspect in terms of currency or content.

7.3 Practical Considerations

The process for selecting V&V tasks rationally within a constrained budget environment revolves around the need for M&S and data credibility balanced by concern for the cost of the V&V activities that contribute to it. Also, while VV&A enhances a simulation's credibility it cannot guarantee that the M&S results will be correct, that the results will be correctly analyzed and interpreted, or that the right model was chosen to solve the problem. A manager must be confident of the value added by the VV&A and VV&C process.

The first thing to consider in using M&S in STEP is to conduct an in-depth analysis to define what the M&S is required to do. Before any decisions are made about applying M&S to a given problem, the problem itself must be defined and articulated clearly enough to permit a precise specification of where M&S will play a role in the solution of the problem, and how it will contribute to the solution of the problem.

The next thing is to develop acceptance criteria for candidate M&S. Having defined what M&S will be required to do (functional requirements), it is necessary to determine how well candidate M&S must do those things. The answer to this question comes in the form of two types of acceptance criteria: operational requirements and fidelity requirements.

Operational requirements are "non-analytical" requirements, in the sense that they do not contribute to resolution of program decisions directly. Instead, these requirements define hardware and software compatibility requirements (e.g., the M&S must run on a certain type of workstation under a certain operating system); pre- and post-processing requirements for M&S data (e.g., M&S inputs or outputs must be converted to special file formats); operations and training support requirements (e.g., M&S cannot have license agreement or operator training requirements because there is no budget for such items, or no time for training); and so on.

Fidelity requirements are the hardest to define, and consist of a listing of how well required M&S functions must correspond to the "real world" in order for the M&S outputs to be considered acceptable for the purpose at hand. Although it is generally possible to specify the kind of V&V that needs to be done to support a given level of credibility, the amount of V&V required to establish credibility for a particular application will still be dependent on a clear definition of how program decisions are affected by M&S outputs.

The reader is referred to the DoD *VV&A Recommended Practices Guide* to read about the remaining steps in the VV&A process.

Ongoing VV&A activities are the price one should expect to pay for ascertaining and maintaining the credibility of the M&S tools. When a simulation is modified, it is usually done with the intent of improving its operation, simulation accuracy or simulation scope. These changes may affect the simulation's suitability for particular applications. The changes must be compared against the user's intent (verification), and the impact of the changes on simulation output must also be compared against the real world system or process to measure the increase or decrease in fidelity (validation). Additionally when the real world changes, or the M&S is used for a purpose different than originally intended, previous VV&A results should be reviewed to determine the impact of these changes on the credibility of the simulation. Since the real world is rarely static over any length of time, it is useful to periodically review the VV&A status of an M&S to ensure consistency with the current projection of the real world.

This page intentionally left blank.

Appendix A - Acronyms and Definitions

Acronyms

ADS Advanced Distributed Simulation

AOA Analysis Of Alternatives

CE Concept Exploration

CFD Computational Fluid Dynamics

COIC Critical Operational Issues and Criteria

CTP Critical Technical Parameter

DIS Distributed Interactive Simulations

DMSO Defense Modeling and Simulation Office

DoD Department of Defense

DSMC Defense Systems Management College

DT Developmental Test; Developmental Testing

DT&E Developmental T&E

EMD Engineering and Manufacturing Development

EOA Early Operational Assessments

HITL Human-In-The-Loop

HLA High Level Architecture

HWIL Hardware-in-The-Loop

JMASS Joint M&S System

LFT&E Live Fire T&E

LRIP Low-Rate Initial Production

MAIS Major Automated Information System

M&S Modeling and Simulation

MDAP Major Defense Acquisition Programs

MEL Master Environmental Library

MNS Mission Need Statement

MOE Measures Of Effectiveness

MOO Measures Of Outcome

MOP Measures Of Performance

MOS Measures Of Suitability

MRTFB Major Range and Test Facilities Base

MSETTM&S Educational Training Tool

MSOSAM&S Operational Support Activity

OA Operational Assessments

ORD Operational Requirements Document

OT Operational Test; Operational Testing

OT&E Operational T&E

SIL System/Software Integration Laboratory

SME Subject Matter Expert

STEP Simulation, Test, and Evaluation Process

SWIL Software-In-The-Loop

T&E Test and Evaluation

TEMP Test and Evaluation Master Plan

TPP Technical Performance Parameters

V&V Verification and Validation

VTTR Virtual Test and Training Range

VV&A Verification, Validation, and Accreditation

VV&C Verification, Validation, and Certification

wwwworld wide web

Definitions

Accreditation. The official certification that a model or simulation is acceptable for use for a specific purpose. [DMSO Glossary of M&S Terms; DoDD 5000.59; M&S Master Plan]

Accuracy. The degree of exactness of a model or simulation, high accuracy implying low error. [DMSO Glossary of M&S Terms, DIS Glossary]

Computer Simulation. A dynamic representation of a model, often involving some combination of executing code, control/display interface hardware, and interfaces to real-world equipment. [DMSO Glossary of M&S Terms]

Distributed Interactive Simulation (DIS). (1) Program to electronically link organizations operating in the four domains; advanced concepts and requirements; military operations; research, development, and acquisition; and training. (2) A synthetic environment within which humans may interact through simulation(s) at multiple sites networked using compliant architecture, modeling, protocols, standards, and data bases. [DMSO Glossary of M&S Terms; M&S Master Plan]

Fidelity. (1) The similarity, both physical and functional, between the simulation and that which it simulates. (2) A measure of the realism of the simulation. (3) The degree to which the representation within a simulation is similar to a real world object, feature, or condition in a measurable or perceivable manner. [DMSO Glossary of M&S Terms, DIS Glossary]

High Level Architecture (HLA). Major functional elements, interfaces, and design rules, pertaining as feasible to all DoD simulation applications, and providing common framework within which specific system architectures can be defined. [DMSO Glossary of M&S Terms]

Human-In-The-Loop (HITL). A model that requires human participation. [DMSO Glossary of M&S Terms; DIS Glossary].

Model. A physics, mathematical, or otherwise logical representation of a system, entity, phenomenon, or process. [DMSO Glossary of M&S Terms; DIS Glossary; M&S Master Plan]

Modeling and Simulation (M&S). The use of models, including emulators, prototypes, simulators, and stimulators, either statically or over time, to develop data as a basis for making managerial or technical decisions. The terms "modeling" and "simulation" are often used interchangeable. [DMSO Glossary of M&S Terms]

Model-Test-Model. An integrated approach to using M&S in support of pre-test analysis and planning; conducting the actual test and collecting data; and post-test analysis of test results along with further validation of models using the test data. [DMSO Glossary of M&S Terms; DSMC Guidebook on M&S]

Simulate. To represent a system by a model that behaves or operates like the system. [DMSO Glossary of M&S Terms; DIS Glossary]

Simulation. A method for implementing a model over time. [DMSO Glossary of M&S Terms; DoDD 5000.59]

a.Live Simulation. A simulation involving real people operating real systems.

b.Virtual Simulation. A simulation involving real people operating simulated systems. Virtual simulations inject HITL in a central role by exercising motor control skills (e.g., flying an airplane), decision skills (e.g., committing fire control resources to action), or communications skills (e.g., as members of a C^4I team).

c.Constructive Model or Simulation. Models or simulations that involve real people making inputs into a simulation that carries out those inputs by simulated people operating simulated systems.[DMSO Glossary of M&S Terms; M&S Master Plan]

Simulator. (1) A device, computer program, or system that performs a simulation. (2) For training, a device which duplicates the essential features of a task situation and provides for direct practice. (3) For DIS, a physical model or simulation of a weapons system, set of weapons systems, or piece of equipment which represented some major aspects of the equipment's operation. [DMSO Glossary of M&S Terms; DIS Glossary]

Stimulate. To provide input to a system in order to observe or evaluate the system's response. [DMSO Glossary of M&S Terms; DIS Glossary]

Synthetic Environments. Internetted simulations that represent activities at a high level of realism from simulations of theaters of war to factories and manufacturing processes. These environments may be created within a single computer or a vast distributed network connected by local and wide area networks and augmented by super-realistic special effects and accurate behavioral models. They allow complete visualization of and total immersion into the environment being simulated. [DMSO Glossary of M&S Terms; M&S Master Plan]

Validation. The process of determining the degree to which a model or simulation is an accurate representation of the real world from the perspective of the intended uses of the model or simulation. [DMSO Glossary of M&S Terms; DoDD 5000.59]

Verification. The process of determining that model or simulation implementation accurately represents the developer's conceptual description and specification. Verification also evaluates the extent to which the model or simulation has been developed using sound and established software engineering techniques. [DMSO Glossary of M&S Terms; DoDD 5000.59]

Virtual Prototype. A computer-based simulation of systems and subsystems which exhibits both geometric and functional realism. This three-dimensional virtual mockup may be used to evaluate prototypes or concepts and provides a common platform which all functional disciplines (design, test, manufacturing, logistics, training, operations, etc.) can work. [DMSO Glossary of M&S Terms; DSMC Guidebook on M&S]

Virtual Prototyping. The process of using a virtual prototype, in lieu of a physical prototype, for test and evaluation of specific characteristics of a candidate design. [DSMC *Virtual Prototyping*]

Appendix B - References

DoD Directive 5000.1, *Defense Acquisition*, March 15, 1996

DoD 5000.2-R, *Mandatory Procedures for Major Defense Acquisition Programs and Major Automated Information Systems Acquisition Programs*, March 15, 1996

DoD Directive 5000.59, *DoD Modeling and Simulation Management*, January 4, 1994

DoD 5000.59-P, *Modeling and Simulation Master Plan*, October 1995 http://www/dmso.mil

E-mail for DoD Component M&S focal point mstrplan@msis.dmso.mil

DoD Instruction 5000.61, *DoD Modeling and Simulation Verification, Validation, and Accreditation*, April 29, 1996

DMSO, *Verification, Validation and Accreditation (VV&A) Recommended Practices Guide*, November 1996

DMSO, Glossary of M&S Terms (Draft), November 3, 1995

http://www.dmso.mil/docslib/mspolicy/glossary/glossary.html

DSMC, *Guidebook on M&S*, September 1994

DSMC, *The Systems Acquisition Managers Guide for the Use of Models and Simulations*

DSMC, *Virtual Prototyping, Concept to Production*

DIS Glossary

M&S Educational Training Tool (MSETT) Glossary

A Description of the DoD Test and Evaluation Process for Electronic Warfare Systems, 13 June 1994